

# Field Methods And Data Processing Techniques Associated With Mapped Inventory Plots'

William A. Bechtold<sup>2</sup>  
Stanley J. Zarnoch<sup>3</sup>

**Abstract**—The U.S. Forest Inventory and Analysis (FIA) and Forest Health Monitoring (FHM) programs utilize a fixed-area mapped-plot design as the national standard for extensive forest inventories. The mapped-plot design is explained, as well as the rationale for its selection as the national standard. Ratio-of-means estimators are presented as a method to process data from mapped inventory plots.

It is particularly timely that we have convened to develop a sampling framework for forest ecosystems that is unified across the entire North American continent. In the United States we have been engaged in a similar effort to standardize our inventory system for the past decade. The purpose of this paper is to describe the sampling design implemented nationally in the U.S., provide background information about how this design evolved, explain how the resulting data can be processed, and describe our experience with this design since its implementation in 1991.

## U.S. National Plot Design

The plot design currently utilized for forest inventory and monitoring in all regions of the U.S. is based on a cluster of four points spaced 36.6 m (120 ft) apart (Figure 1). Each plot consists of a central point surrounded by three satellite points separated by 120 degrees. Each of the four points is circled by a 7.3 m (24.0 ft) fixed-radius subplot, where trees 12.7 cm (5.0 inches) in diameter and larger are measured. Together, all four subplots cover  $\frac{1}{15}$  hectare ( $\frac{1}{6}$  acre).

Each subplot includes a 2.1 m (6.8 ft) fixed-radius microplot used to measure seedlings and saplings smaller than 12.7 cm (5.0 inches) in diameter. Together, all four microplots encompass about  $\frac{1}{188}$  hectare ( $\frac{1}{75}$  acre). Microplots are offset from subplot centers to reduce the impact of trampling by field crews.

Each subplot is surrounded by an annular plot. The annular plot is a ring extending from the edge of each subplot to a total distance of 18 m (58.9 ft) from each subplot center.

The annular plot is optional, designed to sample large trees with diameters of 100 cm (40 inches) or greater in regions characterized by large trees. Annular plots are currently considered for implementation only in the Pacific Northwest region of the U.S.

In addition to the trees measured on these plots, data are also gathered about the area, or setting in which the trees are located. Area data are particularly useful for subdivision of the forest into meaningful categories, and to enable post-stratification. Some of these area attributes are measured (e.g., percent slope), some are assigned by definition (e.g., disturbance history), and some are computed from the tree data (e.g., percent stocking).

For proper analysis it is crucial that the tree data recorded on these plots are properly associated with the area data. To accomplish this, all plots are mapped by "condition class." Survey crews assign an arbitrary number (usually 1) to the first condition class encountered on a plot. This number is then defined by a series of predetermined discrete variables

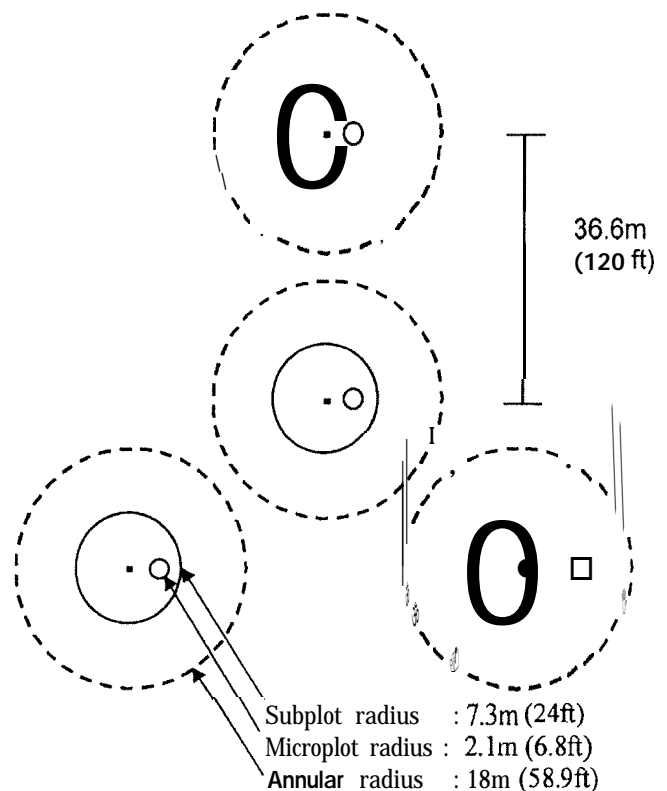


Figure 1 .-Plot configuration.

<sup>1</sup>Paper presented at the North American Science Symposium: Toward a Unified Framework for Inventorying and Monitoring Forest Ecosystem Resources, Guadalajara, Mexico, November 1-6, 1998.

<sup>2</sup>William A. Bechtold is a Research Forester, and USDA Forest Service, Southern Research Station, 200 Weaver Blvd., Asheville, NC, 28802. e-mail: wbechtold/srs,fia@fs.fed.us

<sup>3</sup>Stanley J. Zarnoch is a Mathematical Statistician, USDA Forest Service, Southern Research Station, 200 Weaver Blvd., Asheville, NC, 28802. e-mail: szarnoch/srs,fia@fs.fed.us

attached to it, such as land use, forest type, stand origin, and disturbance history. Additional condition classes are identified if there is a distinct change in any of the condition-class variables on the plot.

It sometimes happens that a single plot straddles two or more distinct conditions. Boundaries between these conditions can bisect the subplots, or they could be located between the subplots. When condition-class boundaries bisect a subplot, they are mapped with a series of two or three azimuths. The crew identifies the condition class at subplot center, and the condition class that contrasts with the condition at subplot center. Standing at subplot center, facing the contrasting condition, they then record the two azimuths where the boundary crosses the subplot perimeter. A third azimuth is permissible if the boundary contains a sharp curve or a corner. Trees tallied on the subplot are then assigned to the condition class in which they occur. The area in each condition is computed when the data are processed. Microplots and annular plots are mapped in a similar fashion. So for each plot the microplot, subplot, and annular area in each condition class is known, as well as the location and condition-class of every tree tallied.

## Rationale for Main Features of the U.S. Plot Design

Many aspects of this design were assembled from a variety of earlier regional designs. Other aspects were contrived to correct known problems with the regional designs. Three important features of the new design deserve further elaboration.

First, it involves clusters of points, as opposed to a single point. Cluster sampling spreads each plot over a larger area than a single-point plot of the same size. This results in larger within-plot variance, but smaller between-plot variance—so population attributes computed from clusters require fewer plots to produce a given sampling error. Field crews thus spend more time collecting data and less time traveling between plots. Another advantage is that clusters are easier to manage in the field. Single-point plots with large radii are more prone to missed trees and other measurement error than the smaller radii of cluster subplots. This is especially true on steep slopes, where slope distance must be converted to horizontal distance. Since the cluster design is more cost-effective for extensive inventories than single-point plots of similar size (Scott 1993), it was commonly utilized in all of our earlier regional inventories, and is now being used for the national design as well.

The second feature of our national design, mapping plots by condition class, is relatively new. This was adopted to overcome problems with the correct matching of tree data with area data. Some of the earlier regional designs addressed this problem by moving plots to confine them to a single condition class. This introduced bias by altering tree selection probabilities. Trees on borders between conditions were clearly under-sampled. Other regional designs addressed this problem by prohibiting the movement of plots, but then blended area data from distinctly different conditions. This was unbiased, but resulted in area misclassifications. For example, a plot that straddled a pure oak forest type and a

pure pine forest type might be classified as a mixed oak/pine forest type. The current design circumvents these problems by prohibiting the movement of plots, and then mapping them to correctly proportion plot area by condition-class variables important for proper stratification of the forest (Hahn et al. 1995).

The third main feature is that the new design is based on fixed-area plots. Prior to implementation of the current design, most regional inventories utilized variable-radius prism plots. The decision to use fixed-radius plots is partially related to difficulties in mapping variable-radius plots. As mentioned previously, some stand-level area attributes (e.g., percent stocking or stand size) are computed from the tree data. For subplots with boundaries this means that tree selection probabilities for some small trees on fragmented subplots could be zero (for the purpose of computing a stand-level attribute for a condition class). Fixed-radius plots avoid this problem since all trees have equal selection probabilities. Additional advantages of fixed-area plots include their utility for spatial modeling, and their intuitive simplicity as compared to variable-radius plots.

## Processing Mapped-Plot Data With Ratios of Means

There are numerous attributes of interest in a forest inventory, but most can be collapsed into three main categories:

1. Per-hectare attributes (e.g., mean number of trees/hectare)
2. Per-tree attributes (e.g., mean number of conks per tree)
3. Stand-level attributes (e.g., mean stand age).

All of these categories involve ratios ( $y/x$ ) where  $y$  is the sum of some attribute of interest and  $x$  is a correlated auxiliary variable (usually an area or a tree total). Population totals, means, and variances associated with these attributes can be derived from several alternative estimators. One such estimator used by our FHM Program, the ratio of means (Cochran 1977), is appealing because it is based on classical statistics and is relatively easy to compute. Demonstrations of how ratios of means apply to mapped inventory data follow. Additional details are covered in Scott and Bechtold (1995) and Zarnoch and Bechtold (1999).

The ratio-of-means estimator is defined as

$$\hat{R} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{\bar{y}}{\bar{x}} \quad (1)$$

where

- $y_i$  = the variable of interest on plot  $i$ ,
- $x_i$  = an auxiliary variable on plot  $i$  which is correlated with  $y_i$  and
- $n$  = number of plots containing at least 1 condition class of interest.

The variance associated with this estimator is

$$\hat{V}(\hat{R}) = \frac{(S_y^2 + \hat{R}^2 S_x^2 - 2\hat{R} S_{yx})}{n(\bar{x}^2)} \quad (2)$$

where  $S_y^2$  and  $S_x^2$  are the typical sample variances of  $y$  and  $x$ , and  $S_{yx}$  is their covariance.

Ratios of means are straightforward for estimates derived singly from only subplots or only microplots. However, simple ratios must be combined when data are aggregated across microplots and subplots. There are two ways to do this. The first method is for attributes where simple ratios can be added together (e.g., estimates of microplot sapling basal area per hectare can be added to subplot basal area per hectare to yield a total for all trees). The additive combined estimator is specified as

$$\hat{R}^* = \hat{R} + \hat{R} \quad (3)$$

where

$\hat{R} =$  the first ratio-of-means estimator, and  
 $\hat{R} =$  the second ratio-of-means estimator.

The variance associated with this estimator, which incorporates the covariance between both estimators, is

$$\hat{V}(\hat{R}^*) = \hat{V}(\hat{R}) + \hat{V}(\hat{R}) + 2\hat{C}(\hat{R}, \hat{R}) \quad (4)$$

where the covariance is defined as

$$\hat{C}(\hat{R}, \hat{R}) = \frac{S_{yy'} - \hat{R}S_{y'x} - \hat{R}'S_{yx} + \hat{R}\hat{R}'S_{xx'}}{n(\bar{x})\bar{x}} \quad (5)$$

A different combined estimator is used for aggregating microplot and subplot ratios that are not additive (e.g., mean number of conks per tree for microplot saplings cannot be added to a similar mean for subplot trees). Nonadditive ratios can be combined by using an estimator that is the ratio of two correlated ratios. This estimator is defined as

$$\hat{R}^{**} = \frac{\hat{R}_1^*}{\hat{R}_2^*} \quad (6)$$

with estimated variance

$$\hat{V}(\hat{R}^{**}) = \hat{R}^{**2} \left[ \frac{\hat{V}(\hat{R}_1^*)}{\hat{R}_1^{*2}} + \frac{\hat{V}(\hat{R}_2^*)}{\hat{R}_2^{*2}} - \frac{2\hat{C}(\hat{R}_1^*, \hat{R}_2^*)}{\hat{R}_1^* \hat{R}_2^*} \right] \quad (7)$$

where

$$\hat{C}(\hat{R}_1^*, \hat{R}_2^*) = \hat{C}(\hat{R}_1, \hat{R}_2) + \hat{C}(\hat{R}_1, \hat{R}_2) + \hat{C}(\hat{R}_1, \hat{R}_2) + \hat{C}(\hat{R}_1, \hat{R}_2) \quad (8)$$

If mean number of conks per tree is the attribute of interest, the ratio in the numerator of Equation (6) would be the combined estimate of conks per hectare (which is additive) and the ratio in the denominator would be the combined estimate of trees per hectare (which is additive).

Combined estimators involving microplots and subplots can easily be extended to include annular plots.

## Estimation of Population Means

Using Equation (1), the simple uncombined ratio, we now illustrate how mapped subplot data are processed for each of the 3 main categories of interest (i.e., per-hectare, per-tree, and stand-level attributes). First, all data are summed across the four subplots and stored at the plot level, by condition class. There is usually no interest in variability among subplots in routine summarizations of inventory data, so the sampling unit is the plot.

For a given plot  $i$ ,  $y_i$ , and  $x_i$  in Equation (1) are defined as

$$y_i = \sum_{j=1}^{c_i} I_{ij} y_{ij}$$

and

$$x_i = \sum_{j=1}^{c_i} I_{ij} x_{ij}$$

where

$y_{ij}$  = the value of the  $y$  variable of interest on plot  $i$  and condition class  $j$

$x_{ij}$  = the value of the  $x$  variable of interest on plot  $i$  and condition class  $j$

$c_i$  = the number of condition classes on plot  $i$

$I_{ij}$  = 1 if condition class  $j$  on plot  $i$  is of interest or

= 0 if condition class  $j$  on plot  $i$  is not of interest.

So, for examples in the three main categories of interest,  $y_{ij}$  and  $x_{ij}$  would be defined for each condition class as follows:

(1) per-hectare attributes (e.g., mean number of trees/hectare)

$y_{ij}$  = sum of the number of all trees tallied on plot  $i$  in condition class  $j$

$x_{ij}$  = area of plot  $i$  in condition class  $j$

(2) per-tree attributes (e.g., mean number of conks per tree)

$y_{ij}$  = sum of all conks on tally trees on plot  $i$  in condition class  $j$

$x_{ij}$  = the number of trees tallied on plot  $i$  in condition class  $j$

(3) stand-level attributes (e.g., mean stand age)

$y_{ij}$  = subplot area of plot  $i$  in condition class  $j$  x (stand age of plot  $i$  in condition class  $j$ )

$x_{ij}$  = subplot area of plot  $i$  in condition class  $j$ .

The condition-class values  $y_{ij}$  and  $x_{ij}$  are thus summed across the condition classes of interest to yield the plot values  $y_i$  and  $x_i$ , which are then summed across plots to produce the ratio of interest.

## Estimation of Population Totals

In the U.S., population totals are usually calculated by expanding per-hectare values by an estimate of total forest area in the region of interest. For example, the mean number of trees per hectare ( $\hat{R}$ ) is multiplied by the total forest area in the region ( $A_F$ ) to yield  $\hat{T}$ , the estimate of total trees in the region:

$$\hat{T} = A_F \hat{R} \quad (9)$$

where

$A_F$  = total forest area in the region

$$\hat{R} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_{Fi}}$$

$y_i$  = number of trees tallied on plot  $i$

$x_{Fi}$  = sum of forest area sampled on plot  $i$ , and

$n$  = the number of plots in the region with at least 1 forested condition class.

Note that the numerator of  $\hat{R}$  in this example is based on the sum of the tree attribute of interest in the condition classes of interest as previously explained, but the denominator is based on the sum of all condition classes in forest. When expanding per-hectare values, the denominator of  $\hat{R}$  must

be formulated on the same basis as the expander ( $A_F$ ). So the condition classes of interest in the numerator of may or may not be the same as the denominator depending on the estimation problem at hand.

Area totals are obtained in the same manner as tree totals. If the population total of interest ( $T$ ) were defined as the total area of a particular forest type,  $y_i$  in the above example would be replaced with the sum of the sampled area in the forest type of interest.

As opposed to expanding on the basis of forest area, total land area could be used as the basis for expansion. If total land area is used,  $A_F$  and  $x_{Fi}$  in equation (9) would be replaced by  $A_T$  and  $x_{Ti}$ , respectively, as follows:

$$\hat{T} = A_T \hat{R} \quad (10)$$

where

$A_T$  = total land area in the region

$$\hat{R} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_{Ti}}$$

$y_i$  = sum of the attribute of interest on plot  $i$

$x_{Ti}$  = sum of total land area sampled on plot  $i$

$n$  = the total number of plots in the region.

The variance associated with the population total ( $\hat{T}$ ) depends on how the expansion factor ( $\hat{A}$ ) is obtained. If the expansion factor is known, as may be the case for total land area in the region of interest, then the sample variance is

$$\hat{V}(\hat{T}) = \hat{A}^2 \hat{V}(\hat{R})$$

If the expansion factor is estimated, as it is in the U.S., where forest area independently estimated from aerial photography is the basis for expansion, then the variance is increased to account for this additional independent estimator

$$\hat{V}(\hat{T}) = \hat{A}^2 \hat{V}(\hat{R}) + \hat{R}^2 \hat{V}(\hat{A})$$

## Implementation of Mapped Inventory Plots

In addition to the U.S., this plot design has been implemented in Lithuania, Latvia, Estonia, Belarus, and Indonesia. No internationally implemented plot design can be expected to achieve optimal efficiency for all inventory attributes in all regions, but so far this design has worked reasonably well wherever implemented. Some criticisms and potential disadvantages have surfaced, but all are minor and solutions are available to circumvent them. A few remarks regarding concerns about this design should be of interest.

Total numbers of trees tallied on each plot average around 30-60, depending on the region. Once a crew reaches the site, most plots are completed in about half a day. However, some regions with high populations of planted trees in the 13-18 cm (5-7 inch) diameter range have complained that too many trees are sampled per plot. Variance between plantation trees is fairly low, and a  $\frac{1}{15}$  ha ( $\frac{1}{6}$  acre) fixed plot in a young plantation often yields upwards of 100 tally trees. For regions with a high proportion

of young planted stands, sampling efficiencies might be improved by dropping one of the four subplots to reduce the number of tally trees. Although all regions are currently required to measure all four subplots, some consideration should be given to making the fourth subplot optional.

There has been some concern that mapped designs might yield unrealistic values for conditions occupying a small fragment of a plot. This is definitely not a problem for population estimates because the ratio-of-means estimator pools all conditions of interest prior to computing population means or totals. On the other hand, such fragments might pose problems for some discrete area classifications computed from tree tally, such as forest type. Insufficient tree tally on small fragments can be handled with supplemental sampling for the purpose of area classification, or reliance on subjective classification by field crews.

Some indices, such as species diversity index, work best when all plots are the same size. Multiple condition classes on mapped plots can always be collapsed back together to satisfy this requirement. This is the case with the U.S. design, except where plots contain mixtures of forest and nonforest, because no vegetation is sampled on nonforest lands. Even so, species-area curves (Pielou 1966) could be used to adjust partially forested plots, or partially forested plots could simply be excluded from some analyses, or we could simply decide to sample nonforest vegetation on plots that contain mixtures of forest and nonforest.

On the surface it appears mapped designs have the potential to create an unwieldy number of condition-class permutations at the regional level, or even at the plot level. When the data are processed with ratios of means, this is clearly not the case. For any attribute of interest, all condition classes eventually collapse into two—those condition classes that are of interest, and those that are not.

All things considered, the fixed-area mapped-plot design has proven to be robust across a wide variety of field conditions. It has some minor disadvantages, as would any broad-scale design implemented across such a wide range of field conditions, but these are readily overcome. This design now forms the basis of a unified sampling framework across the U.S. It also deserves serious consideration as the choice for implementing a unified sampling framework across the North American continent.

## Literature Cited

- Cochran, W. G. 1977. Sampling techniques. Third Edition. John Wiley & Sons, Inc. New York. 428 p.
- Hahn, J. T., C. D. MacLean, S. L. Arner and W. A. Bechtold. 1995. Procedures to handle FIA cluster plots that straddle two or more conditions. For. Sci. Monogr. 31:12-25.
- Pielou, E.C. 1969. An introduction to mathematical ecology. John Wiley & Sons, Inc. New York. 286 p.
- Scott, C. T. and W. A. Bechtold. 1995. Optimal design of a plot cluster for monitoring. In: The optimal design of forest experiments and forest surveys, Sept 10-14. School of Math, Statistics and Computing. Univ. of Greenwich, London. Pages. 233-242.
- Scott, C. T. and W. A. Bechtold. 1995. Techniques and computations for mapping plot clusters that straddle stand boundaries. For. Sci. Monogr. 31:46-61.
- Zarnoch, S.J., and W.A. Bechtold. 1999. Estimating mapped-plot attributes with ratio of means. For. Sci. In review.